Section I. Data sources

Data sources are given in Table A1.

Table A1. Data sources

No	Dataset	Full name	Website
1	ESDB	European Soil Database	http://eusoils.jrc.ec.europa.eu/esdb_archive/ESDB/Index.htm
2	GSM	General Soil Map of United States	http://soildatamart.nrcs.usda.gov/USDGSM.aspx
3	SLC	Soil landscapes of Canada	http://sis.agr.gc.ca/cansis/nsdb/slc/index.html
4	China	The soil database of China for land surface modeling	http://globalchange.bnu.edu.cn/research/soil2
5	ASRIS	the ASRIS (Australian Soil Resource Information System) polygon attributed surfaces	http://adl.brs.gov.au/anrdl/php/anrdlSearch.html
6	SOTWIS	soil property estimates derived from the WISE and SOTER (Soil Terrain Database)	http://www.isric.org/data/data-download
7	DSMW	Digital Soil Map of the World	http://www.fao.org/nr/land/soils/digital-soil-map-of-the-world/en/
8	WISE	World Inventory of Soil Emission Potential profile database	http://www.isric.org/data/isric-wise-global-soil-profile-data-ver-31
9	NCSS	National Cooperative Soil Survey of United States profile database	http://soils.usda.gov/contact/nssc/

1. European Soil Database (ESDB)

The ESDB [ESB, 2004] contains 4 components: the Soil Geographical Database of Eurasia at scale 1:1,000,000 (SGDBE), the Pedotransfer Rules Database (PTRDB), the Soil Profile Analytical Database of Europa (SPADBE), and the Database of Hydraulic Properties of European Soils (HYPRES). In this study, we used SGDBE4 and SPADBE2.0. SGDBE was compiled in the 1970's but considerably updated in the 1990s. SGDBE contains a list of Soil Typological Units (STU). Besides the soil names they represent, these units are described by variables (attributes) specifying the nature and properties of the soils: for example the texture, the water regime, the stoniness, etc. STUs are grouped into Soil Mapping Units (SMU) to form soil associations and to illustrate the functioning of pedological systems within the landscapes. SPADE2 was developed to derive appropriate characterization of soil profile data for STUs in the SGDBE. SPADE 2 aims to provide sufficient soil property data to support higher tier modeling of pesticide fate at the European level. However, it only covers limited numbers of countries in Europe, including Belgium and Luxembourg, Denmark, England and Wales, Finland, Germany, Italy, Netherlands, Portugal, Scotland Bulgaria, Estonia, France, Hungary, Ireland, Romania, Slovakia and Switzerland.

2. General Soil Map (GSM) of U.S.

GSM was developed by the National Cooperative Soil Survey and supersedes the State Soil Geographic (STATSGO) data set published in 1994[*USDA-NCSS*, 2006]. The data set was created by generalizing more detailed soil survey maps. Data on geology, topography, vegetation, climate and Land Remote Sensing Satellite (LANDSAT) images were assembled, where more detailed soil survey maps were not available. Soils in similar conditions were studied, and the probable classification and extent of the soils were determined. Map unit composition was determined by transecting or sampling areas on the more detailed maps and expanding the data statistically to characterize the whole map unit. This data set consists of spatial vector data and tabular data. Soil map units are linked to attributes in the National Soil Information System data base which gives the proportion of soil components and their properties.

3. Soil landscapes of Canada (SLC)

The SLCs are a series of digital maps that show the major characteristics of soil and land for Canada [*Soil Landscapes of Canada Working Group*, 2010]. SLCs were recompiled at a scale of 1:1 million based on existing soil survey maps. Each map unit is described by a standard set of attributes. The full array of attributes that describe a distinct type of soil and its associated landscape, such as surface form, slope, water table depth, permafrost and lakes, is called a soil landscape. SLC polygons may contain one or more distinct soil landscape components. SLCs were originally conceived as a standardized database consisting of major attributes important to plant growth, land management, and soil degradation. SLC version 3.2 is the latest revision. It provides soil information for the major agricultural regions of Canada, although Alberta, Nova Scotia and Prince Edward Island have component, soil name and soil layer data for the entire province (i.e. beyond the agricultural areas).

4. The soil database of China for land surface modeling

The soil database of China for land surface modeling is a comprehensive 30×30 arc-second resolution gridded soil characteristics dataset [*Shangguan et al.*, 2013]. It includes physical and chemical attributes of soils derived from 8,979 soil profiles and the Soil Map of China (1:1,000,000). The profiles and soil map are from the Second National Soil Survey (1979-1985). There are only 925 soil map units. Unlike most of other soil maps, each map unit has only one component in the soil map. There are 94,303 polygons in the soil map with 85,257 soil polygons. We used the polygon linkage method to derive the spatial distribution of soil properties. The profile attribute database and soil map are linked under the framework of the Genetic Soil Classification of China which avoids uncertainty in taxon referencing. Quality control information is included to provide ‘confidence’ information for the derived soil parameters.

5. ASRIS (Australian Soil Resource Information System) polygon attributed surface

The ASRIS polygon attributed surface modeled from area based observations made by soil agencies both State and CSIRO (Commonwealth Scientific and Industrial Research Organisation) and presented as 0.01 degree grid cells [*CSIRO*, 2001]. The final ASRIS polygon attributed surfaces are a mosaic of all of the data obtained from various state and federal agencies. The surfaces have been constructed with the best available soil survey information available at the time. The surfaces also rely on a number of assumptions. One being that an area weighted mean is a good estimate of the soil attributes for that polygon or map-unit. The polygon data was then converted to a continuous raster surface using the soil attribute values calculated for each polygon. Another assumption made is that the look-up tables provided by [*McKenzie et al.*, 2000], state and territories accurately depict the soil attribute values for each soil type. In cases where a soil type was missing from the look-up table or layer 2 did not exist for that soil type, the percent area of the soils remaining were adjusted prior to calculating the final soil attribute value. The accuracy of the maps is most dependent on the scale of the original polygon data sets and the level of soil survey that has taken place in each state. The Atlas of Australian Soils is considered to be the least accurate dataset and has therefore only been used where there is no state based data. The state datasets, including Western Australian sub-systems, South Australian land systems and NSW soil landscapes and reconnaissance mapping, would be the most reliable based on scale. NSW soil landscapes and reconnaissance mapping use only one dominant soil type per polygon was used in the estimation of attributes. South Australia and Western Australia use several soil types per polygon or map-unit.

6. SOTER derived databases(referred as SOTWIS)

The SOTER (SOil and TERrain database) approach is based on land system to re-inventory global land resources. This approach was implemented in many regions all over the world. Though the information is collected according to the same SOTER methodology, the results of each region are in different scales. The WISE (World Inventory of Soil Emission Potential) is used to fill gaps in measured soil physical and chemical data in primary SOTER databases, resulting in so-called SOTWIS databases [Batjes, 2003; 2007; van Engelen et al., 2005]. SOTWISE contains soil parameter estimates for five standard depths (0–20 cm, 20–40cm, 40–60 cm, 60–80 cm and 80–100cm) and five soil textural classes (coarse, medium, medium fine, fine and very fine). The SOTER databases contain more than 1,800 geo-referenced soil profiles for South and Central America and the Caribbean, more than 900 geo-referenced profiles for Southern Africa, more than 600 geo-referenced profiles for Central and Eastern Europe, some not fully geo-referenced profiles for northeastern Africa, and very limited profiles for north and central Eurasia. In this study, the SOTWIS of following regions were included: Central Africa [*Batjes*, 2007], Indo-Gangetic Plains[Batjes et al., 2004], Jordan [Batjes et al., 2003] and Kenya [Batjes and Gicheru, 2004], Latin America and the Caribbean [*FAO/UNEP/ISRIC/CIP*, 1998], Southern Africa [*FAO/ISRIC*, 2003], Senegal and The Gambia [*Batjes*, 2008a], Tunisia [*Batjes*, 2010].

7. Digital Soil Map of the World (DSMW)

The 1:5 million scale Digital Soil Map of the World (DSMW) is the only world-wide soil map in digital format [*FAO*, 1995, 2003]. It was widely used and almost all the derived global soil datasets were based on this map. This map was compiled using the data between the 1930s and the 1970s. In the digital version, it contains the vector and raster maps with composition of the soil units, top soil texture, slope class and soil phase for more than 5,000 map units, and statistically derived soil properties such as pH, organic matter, C/N, soil moisture storage capacity and soil depth. It also contains interpretations by country on the extent of specific problem soils, the fertility capability classification results by country and corresponding maps.

8. World Inventory of Soil Emission Potential profile database (WISE)

WISE was compiled based on soil profile data collected by soil professionals worldwide [*Batjes*, 2008b]. It includes primary soil data and derived secondary data. Methods and standards to sample, describe and analyze the profiles differ in different countries. The profiles do not have a uniform set of properties, generally because the original survey had selected measurements. Laboratory methods of specific soil properties vary between laboratories and over time. WISE has strict criteria for accepting profiles. Sometimes, results for the same property may not be comparable. The geographic and taxonomic coverage of profiles are uneven because the profile representation is based on the availability of sufficiently detailed legacy data. The gaps in WISE can be of a taxonomic, geographic, and soil analytical nature. As a result, not all the data can be used for application purposes.

9. National Cooperative Soil Survey of United States profile database (NCSS)

The NCSS is developed by the National Soil Survey Laboratory of the National Cooperative Soil Survey [NCSS, 2012]. The availability of different soil characterization data varies because only selected measurements were planned. The database contains completed project information of the Soils Survey Laboratory. For research purposes, pedons are selected to represent the central concept of a soil series or the central concept of a map unit, or to bracket a range of properties within a series or landscape. Analytical procedures and methods of soil preparation are taken from the standard of soil survey laboratory methods manual.

Section II. Three Mapping Approaches

Different soil property mapping approaches gave different soil property estimates for a grid. Here, we take the SOC of layer 2 as an example to show the differences by the three methods (Figure S1 and S20). For the convenience of comparison, we prepared the maps by Method A and Method D in the pre-selected classes of Method B. SOC by Method D was usually lower than SOC by Method A or Method B in some areas of the North Africa and the Near East. This indicates that SOC of the dominant component in a mapping unit was lower than the average SOC of all components in these areas. However, there were some tropic soils in Southeast Asia with a higher SOC by Method D. Method D tended to give more extreme values than the other two, because it only considers the dominant soil mapping unit, while the other two may take more than one map mapping unit compositions into account. In the North America, ESDB regions and SOTWIS regions, the map by Method B had a similar spatial pattern to the map by Method D. This indicates that the value of the dominant soil types belonged to the dominant soil attribute class in most areas. For China, the three methods had identical estimates because there is only one soil type in a map unit there.

Each mapping methods had its advantages and disadvantages [*Batjes*, 2006]. Data users should choose the map derived using the three mapping approaches according to their applications. Method A and Method D provide un-binned values for each grid cell which makes model running convenient, but Method B does not. Method A is the only method which can keep mass conservation for soil properties such as SOC and total nitrogen. However, Method A may mislead in some cases. For example, if a grid cell is comprised of 80% of mineral soil with 2% SOC and 20% of organic soil with 50% SOC, Method A will give an estimation of 11.6% SOC. This will makes the grid cell appear to be an organic soil. For the soil property recorded in a logarithmic scale such as soil pH, Method A will be more misleading. Though Method B only provided binned classes, it is considered more appropriate to represent a grid cell than Method D or A, especially when the percentage of dominant soil components is low.

Section III. General Information Maps

For general information maps, they were derived using Method B and only the dominant class of a map unit was shown.

1. FAO symbols

Figure S2 shows the FAO legends including FAO74, FAO85 and FAO90. The information of FAO74 was from the DSMW and covered worldwide. The information of FAO85 legend was from the ESDB and covered Europe and Russia. The information of FAO90 was from ESDB, SOTWIS and China and covered these areas.

2. Non-soil class

Figure S3 shows the non-soil classes in the FAO legends. 10 classes were identified, i.e. inland water, urban, salt flats, rock debris, no data, island, humanly disturbed, glaciers and permanent snow, fishpond and dunes and shifting sands.

3. topsoil texture class

Figure S4 shows topsoil(0-0.3m) texture classes. Only three simplified textural classes were used. Most soils of the world were medium texture.

4. soil drainage class

Figure S5 shows the soil drainage classes. The six drainage classes were very poor, poor, imperfectly, moderate well, well, somewhat excessive. Drainage classes represent reference drainage conditions assuming flat terrain (i.e., 0.0 - 0.5% slope) [*FAO/IIASA/ISRIC/ISS-CAS/JRC*, 2012]. Most soils of the world were moderate well.

5. reference soil depth

Figure S6 shows the soil reference depth classes. Three arbitrary classes were used, i.e., 10, 30 and 100 cm. These were not the actual soil depth. It should be noted that there are many soils with a depth (far) more than 100cm.

6. available water storage capacity class

Figure S7 shows the available water storage capacity classes. The AWC classes were estimated from FAO legend, topsoil textural class and depth/volume limiting soil phases[*FAO/IIASA/ISRIC/ISS-CAS/JRC*, 2012].

7. soil phase

Figure S8 shows soil phases. Only the first two dominant phases in a map unit were given.

8. obstacles to roots

Figure S9 shows the depth of an obstacle to roots. Only the ESDB had the information.

9. impermeable layer

Figure S10 shows the depth of an impermeable layer. Only the ESDB had the information.

10. soil water regime

Figure S11 shows the dominant annual average soil water regime classes. Only the ESDB had the information. Code 1 indicates not wet within 80 cm for over 3 months, nor wet within 40 cm for over 1 month. Code 2 indicates wet within 80 cm for 3 to 6 months, but not wet within 40 cm for over 1 month. Code 3 indicates wet within 80 cm over 6 months, but not wet within 40 cm for over 11 month. Code 4 indicates wet within 40 cm depth for over 11 month.

11. Additional property

Figure S12 shows the additional property for agriculture use. The three classed were vertic, gelic and petric [*FAO/IIASA/ISRIC/ISS-CAS/JRC*, 2012].

Section IV. Soil Property Maps

For soil property maps, they were derived using Method A and the average value of a map unit was shown.

1. Sand, silt and clay content

Figure S13 shows the geographic distributions of sand, silt and clay content for layer 2 and 6. As the depth increase, there was an increase in area of regions with high clay content and decrease in area of regions with high sand content. It should be noted that some grids such as those in the Greenland may have zero value for sand, silt and clay content, where there was no soil.

2. Gravel content

Figure S14 shows the geographic distributions of gravel content for layer 2 and 6. Most soil of the world had low gravel content. High values were found in the high latitude of the north hemisphere, Qinghai-Tibet Plateau, west part of the US and some areas in the South America. The gravel content decreased in the high latitude of the north hemisphere, but increased in the Qinghai-Tibet Plateau.

3. Bulk density

Figure S15 shows the geographic distributions of BD for layer 2 and 6. High BD appeared in the US for layer 2 and in the US, Russian Far East, India and the east of Australia for layer 6. Soils with high SOC were corrected to a low value of BD in the high latitude of north hemisphere. BD increased with depth in most areas. Like the PSD, some grids may have zero value for BD, where there was no soil.

4. Volumetric water content at -10, -33 and -1500 kPa

Figure S16, S17 and S18 shows the geographic distributions of volumetric water content at -10, -33 and -1500 kPa for layer 2 and 6, respectively. High water content were scattered and low water content were found in desert areas. The changes with depth were limited to small areas.

5. Total carbon and soil organic carbon

Figure S20 shows the geographic distributions of total C for layer 2 and 6. High value of total C appeared in the high latitude of the north hemisphere and low value appears in the desert areas. total C decreased with the depth in most areas.

Figure S20 shows the geographic distributions of SOC for layer 2 and 6. The distribution of SOC was similar to total C. SOC decreased a lot from layer 2 to layer 6 in most areas. However, high SOC still appeared in some areas of the high latitude of north hemisphere.

6. Total nitrogen

Figure S21 shows the geographic distributions of total N for layer 2 and 6. Similar to total C, high value of total N appeared in the high latitude of the north hemisphere and low value appears in the desert areas. Total N decreased a lot from layer 2 to layer 6.

7. Total phosphorus

Figure S22 shows the geographic distributions of total P for layer 2 and 6. Most soils had a low total P and high value of total P appeared in some areas of China and South America. Total P decreased a lot from layer 2 to layer 6.

8. Total potassium

Figure S23 shows the geographic distributions of total K for layer 2 and 6. High total K appeared in the North Africa, Middle East, China and Australia and low total K appeared in scattered areas. There was no significant change with depth.

9. Total sulfur

Figure S24 shows the geographic distributions of total S for layer 2 and 6. High total S appeared in the Middle East, Central Asia and South Africa and low total S were found across the world. Total S decreased with depth in most areas.

10. Calcium carbonate content

Figure S25 shows the geographic distributions of calcium carbonate content for layer 2 and 6. Most soils had low $CaCO_3$ and high $CaCO_3$ appeared from the Middle East to the West China. The change with depth was not obvious.

11. Gypsum content

Figure S26 shows the geographic distributions of gypsum content for layer 2 and 6. Most soils had low gypsum and some small areas in the North Africa and middle Asia had high gypsum. The gypsum content tended to be increased with depth.

12. pH, measured in water, KCL solution and CaCl2 solution

Figure S27, S28 and S29 show the geographic distributions of pH measured in water, KCL solution and $CaCl_2$ solution content for layer 2 and 6, respectively. High pH value distributed mainly from the North Africa to the North China, in the West US and Australia. There were no clear changes with depth. The pH measured in water was usually higher than pH measured in $CaCl_2$, and pH measured in KCl was the lowest.

13. Electrical conductivity

Figure S30 shows the geographic distributions of electrical conductivity for layer 2 and 6. Most soils had a low ECE, and the high value of ECE existed in the middle Asia. And ECE tended to decrease with depth in the areas with high ECE.

14. Exchangeable cations, cation exchange capacity, base saturation and exchangeable sodium percentage

Figure S31-S36 show the geographic distributions of exchangeable cations (H^+, Al^{3+}, Ca^{2+}, Mg^{2+}, K^+, Na^+) for layer 2 and 6. Figure S37 shows the geographic distributions of cation exchange capacity for layer 2 and 6. High CEC appeared in the high latitude of the north hemisphere and low CEC appeared in the Africa, middle Asia and South America. The CEC seems to decrease with depth. Figure S38 shows the geographic distributions of base saturation for layer 2 and 6. Most soils were with high base saturation and low base saturation appeared in the middle Africa, Southeast Asia, north part of South America and East Canada. Figure S39 shows the geographic distributions of exchangeable sodium percentage for layer 2 and 6. Most soil had a low exchangeable sodium percentage, and high value only appears in small areas.

15. phosphorous measured in different method

Figure S40-S44 show the geographic distributions of phosphorous measured in different methods for layer 2 and 6. There was some lack of data because of the source data. Details about the phosphorous measured in different methods were discussed by [*Batjes*, 2011]

REFERENCES

Batjes, N. H. (2003), A taxotransfer rule-based approach for filling gaps in measured soil data in primary SOTER databases, International Soil Reference and Information Centre, Wageningen.

Batjes, N. H., Z. Rawajfih, and R. Al-Adamat (2003), Soil data derived from SOTER for studies of carbon stocks and change in Jordan (ver. 1.0), ISRIC - World Soil Information, Wageningen.

Batjes, N. H., T. Bhattacharyya, C. Mandal, K. Dijkshoorn, D. K. Pal, E. Milne, and K. S. Gajbhiye (2004), Soil data derived from SOTER for studies of carbon stocks and change in the Indo-Gangetic Plains (India) (ver. 1.0; GEFSOC Project). Report 2004/06, ISRIC - World Soil Information, Wageningen and NBSS&LUP (ICAR), India.

Batjes, N. H., and P. Gicheru (2004), Soil data derived from SOTER for studies of carbon stocks and change in Kenya (ver. 1.0), ISRIC - World Soil Information, Wageningen.

Batjes, N. H. (2006), ISRIC-WISE derived soil properties on a 5 by 5 arc-minutes global grid. Report 2006/02, ISRIC- World Soil Information, Wageningen (with data set).

Batjes, N. H. (2007), SOTER-based soil parameter estimates for Central Africa - DR of Congo, Burundi and Rwanda (SOTWIScaf, version 1.0), ISRIC - World Soil Information, Wageningen.

Batjes, N. H. (2008a), SOTER parameter estimates for Senegal and The Gambia derived from SOTER and WISE (SOTWIS-Senegal, version 1.0), ISRIC - World Soil Information, Wageningen.

Batjes, N. H. (2008b), ISRIC-WISE harmonized global soil profile dataset (ver. 3.1), ISRIC - World Soil Information, Wageningen.
Batjes, N. H. (2010), Soil property estimates for Tunisia derived from SOTER and WISE. (SOTWIS-Tunisia, version 1.0), ISRIC - World Soil Information, Wageningen.
Batjes, N. H. (2011), Overview of soil phosphorus data from a large international soil database, Plant Research International (PRI), Wageningen UR, and ISRIC - World Soil Information, Wageningen.
CSIRO (2001), Soil Bulk Density for Australian areas of intensive agriculture of Layer 1 and 2 (derived from soil mapping), CSIRO, Land & Water, Canberra.
ESB (2004), European Soil Bureau European Soil Database (vs. 2.0), edited, European Commission- JRC - Institute for Environment and Sustainability, Ispra, Italy.
FAO (1995, 2003), The Digitized Soil Map of the World Including Derived Soil Properties (version 3.6), FAO, Rome.
FAO/IIASA/ISRIC/ISS-CAS/JRC (2012), Harmonized World Soil Database (version1.2), FAO, Rome, Italy and IIASA, Laxenburg, Austria.
FAO/ISRIC (2003), Soil and Terrain Database for Southern Africa, FAO, Rome.
FAO/UNEP/ISRIC/CIP (1998), Soil and terrain digital database for Latin America and the Caribbean at 1:5 Million scale, FAO, Rome.
McKenzie, N. J., D. W. Jacquier, L. J. Ashton, and H. P. Cresswell (2000), Estimation of Soil Properties Using the Atlas of Australian Soils.Technical Report 11/00, , CSIRO Land and Water, Canberra.
NCSS (2012), National Cooperative Soil Characterization Database, National Cooperative Soil Survey.
Shangguan, W., Y. Dai, B. Liu, A. Zhu, Q. Duan, L. Wu, D. Ji, A. Ye, H. Yuan, Q. Zhang, D. Chen, M. Chen, J. Chu, Y. Dou, J. Guo, H. Li, J. Li, L. Liang, X. Liang, H. Liu, S. Liu, C. Miao, and Y. Zhang (2013), A China Dataset of Soil Properties for Land Surface Modeling, *Journal of Advances in Modeling Earth Systems*, *5*, 212-224.
Soil Landscapes of Canada Working Group (2010), Soil Landscapes of Canada version 3.2., Agriculture and Agri-Food Canada.
USDA-NCSS (2006), Digital General Soil Map of U.S., edited, U.S. Department of Agriculture, Natural Resources Conservation Service, Fort Worth, Texas.
van Engelen, V. W. P., N. H. Batjes, J. A. Dijkshoorn, and J. R. M. Huting (2005), Harmonized Global Soil Resources Database, FAO and ISRIC, Wageningen.

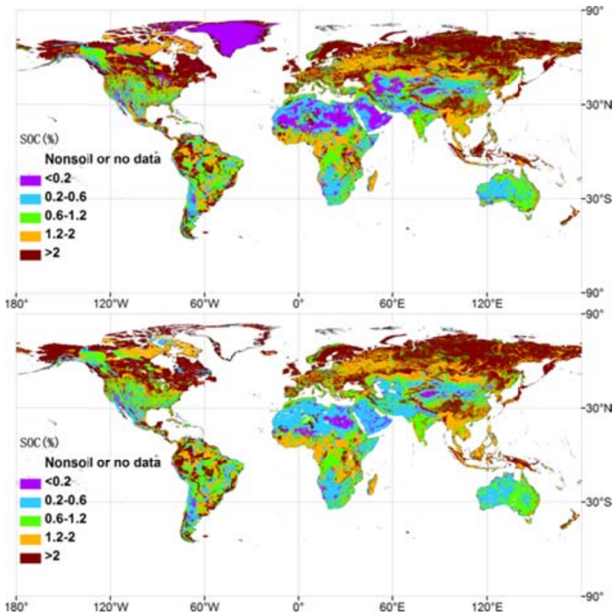

Figure S1. The geographic distribution of soil organic carbon content of layer 2 (0.045-0.091 m) by the dominant soil unit method (Method D, top) and the dominant binned soil attribute method (Method B, bottom).

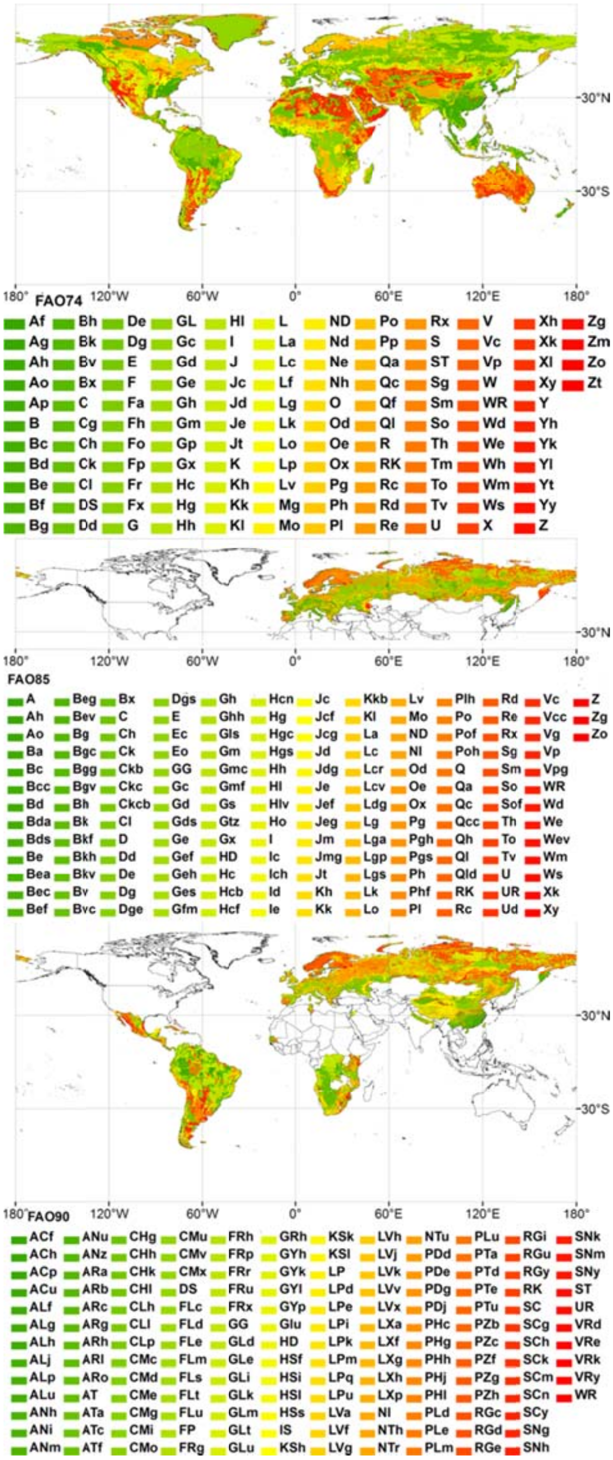

Figure S2. FAO legends.

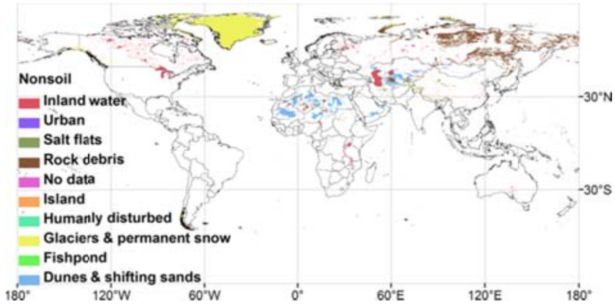

Figure S3. Non-soil classes.

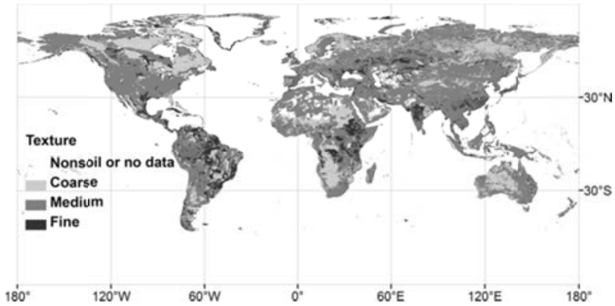

Figure S4. Topsoil texture classes.

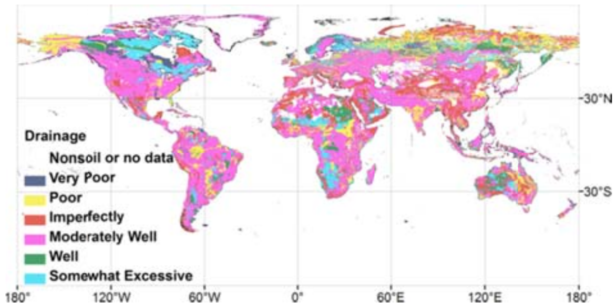

Figure S5. Soil drainage classes.

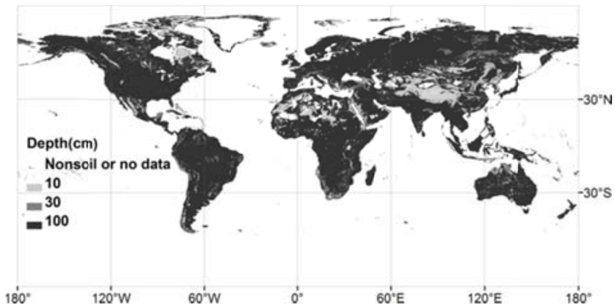

Figure S6. Soil reference depth classes.

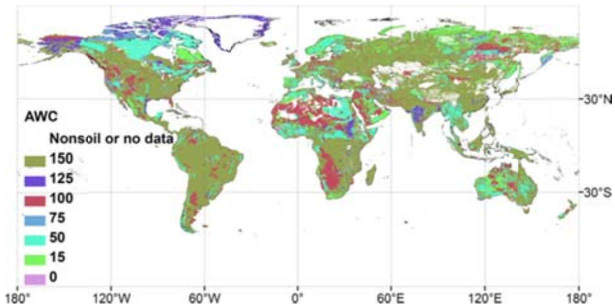

Figure S7. Available water storage capacity classes (mm/m).

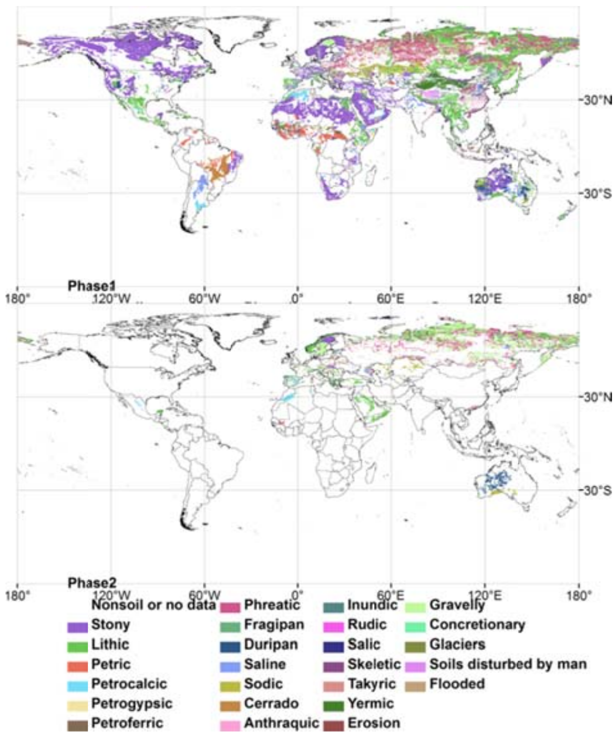

Figure S8. Soil Phases.

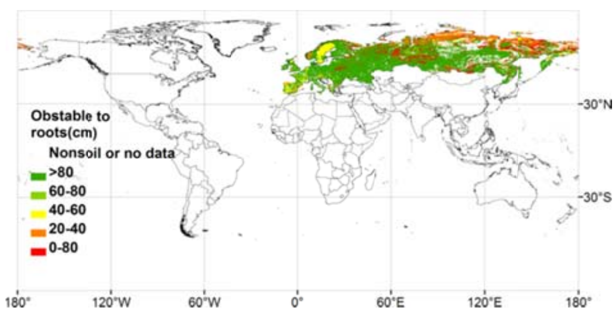

Figure S9. Obstacle to roots.

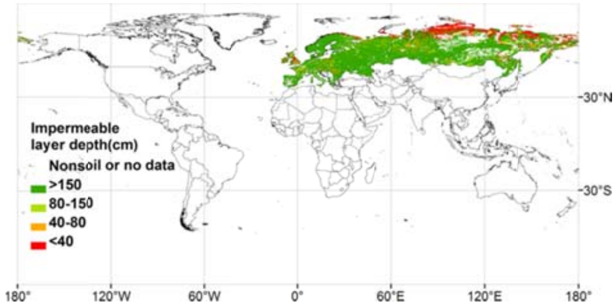

Figure S10. Impermeable layer depth.

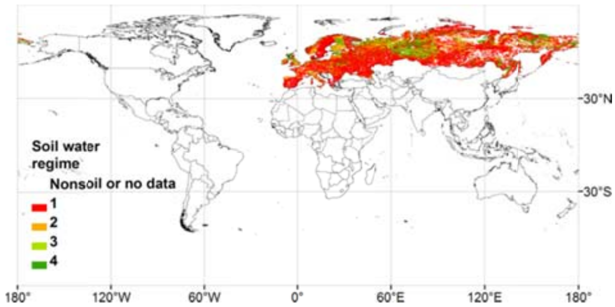

Figure S11. Soil water regime.

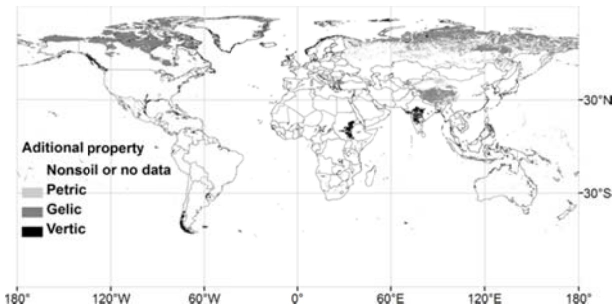

Figure S12. Additional property for agriculture use.

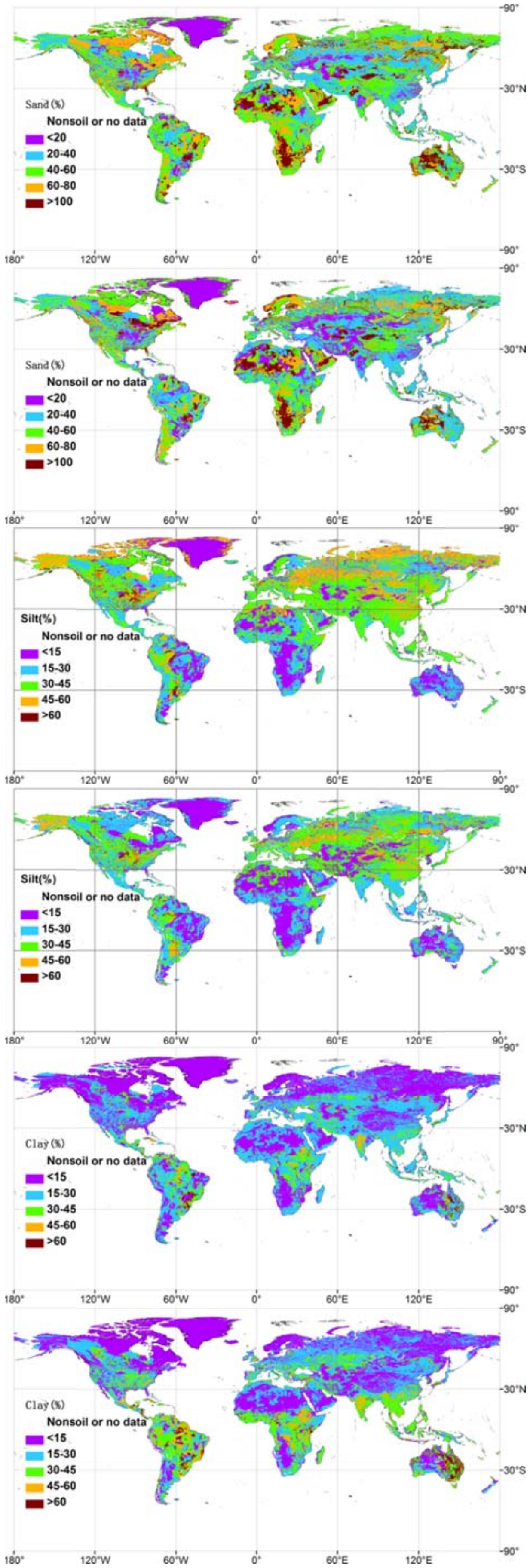
Sand (%)
Nonsoil or no data
<20
20-40
40-60
60-80
>100
Sand (%)
Nonsoil or no data
<20
20-40
40-60
60-80
>100
Silt(%)
Nonsoil or no data
<15
15-30
30-45
45-60
>60
Silt(%)
Nonsoil or no data
<15
15-30
30-45
45-60
>60
Clay (%)
Nonsoil or no data
<15
15-30
30-45
45-60
>60
Clay (%)
Nonsoil or no data
<15
15-30
30-45
45-60
>60
90°
30°N
30°S
180°
120°W
60°W
0°
60°E
120°E

Figure S13. The geographic distribution of sand, silt and clay content (%) of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

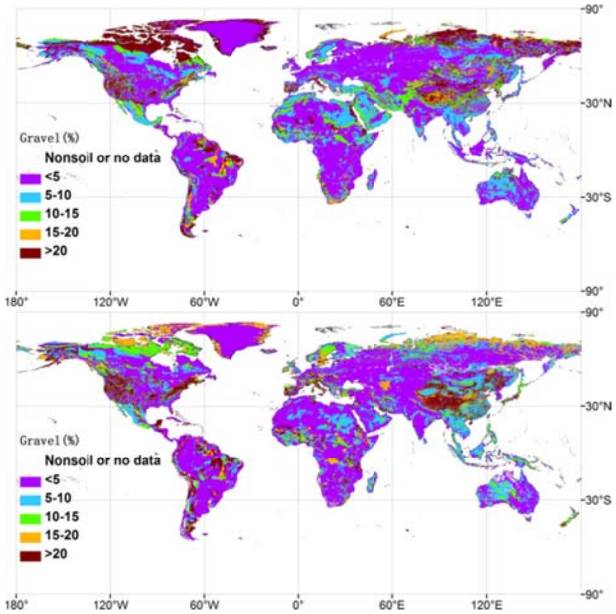

Figure S14. The geographic distribution of gravel content (%) of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

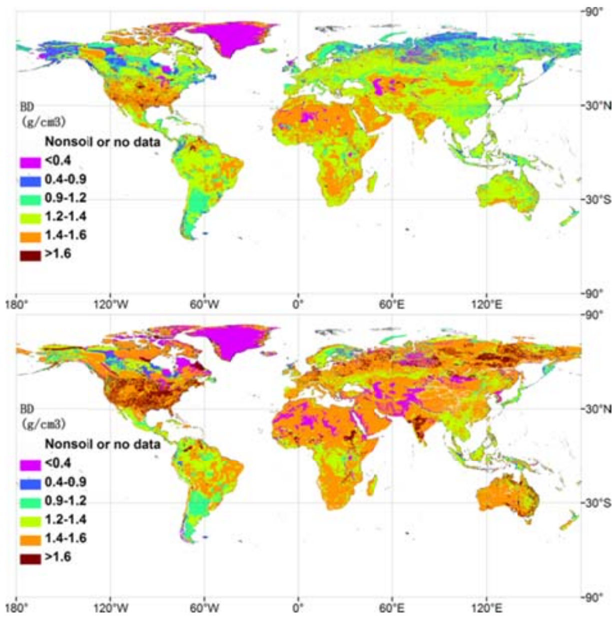

Figure S15. The geographic distribution of bulk density of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

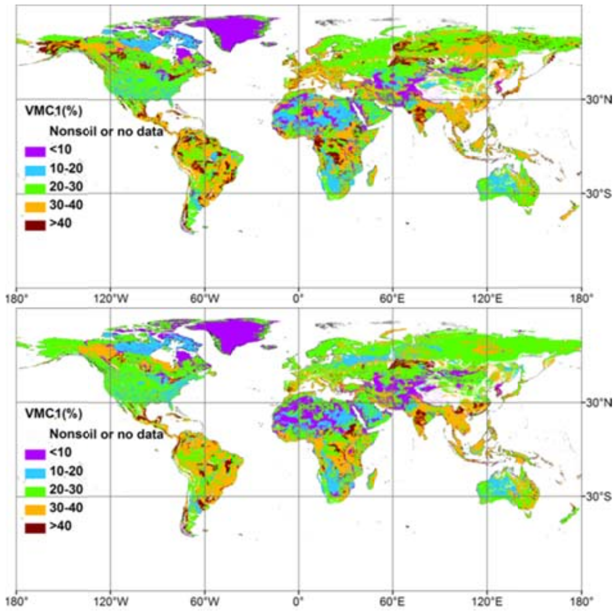

Figure S16. The geographic distribution of volumetric water content at -10 kPa of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

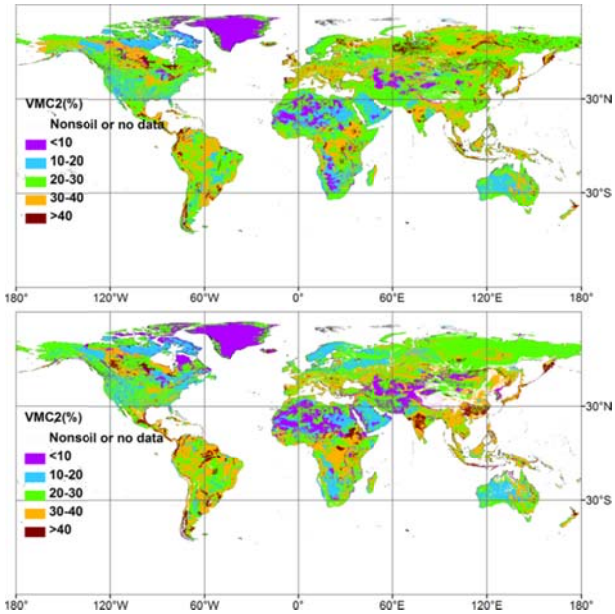

Figure S17. The geographic distribution of volumetric water content at -33 kPa of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

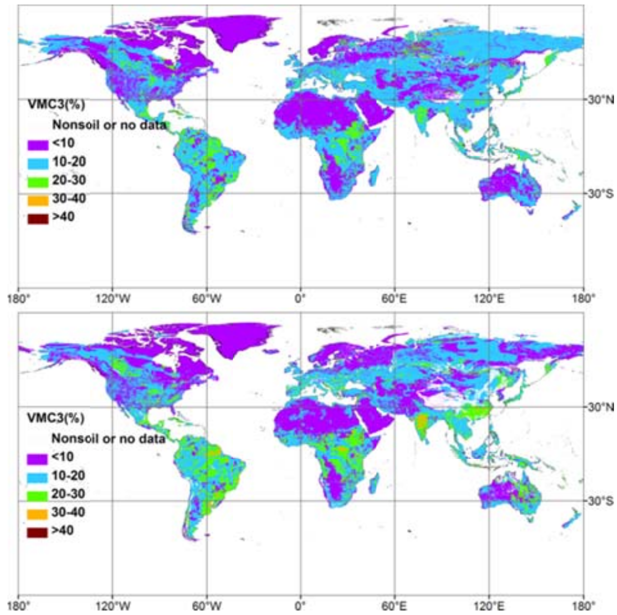

Figure S18. The geographic distribution of volumetric water content at -1500 kPa of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

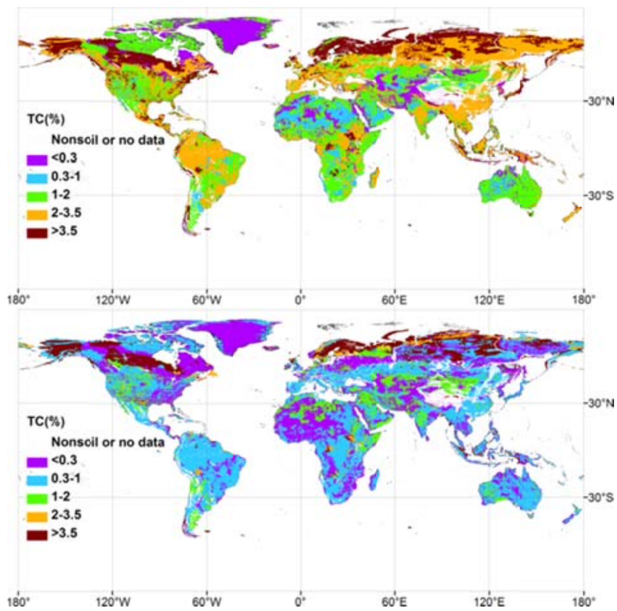

Figure S19. The geographic distribution of total carbon of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

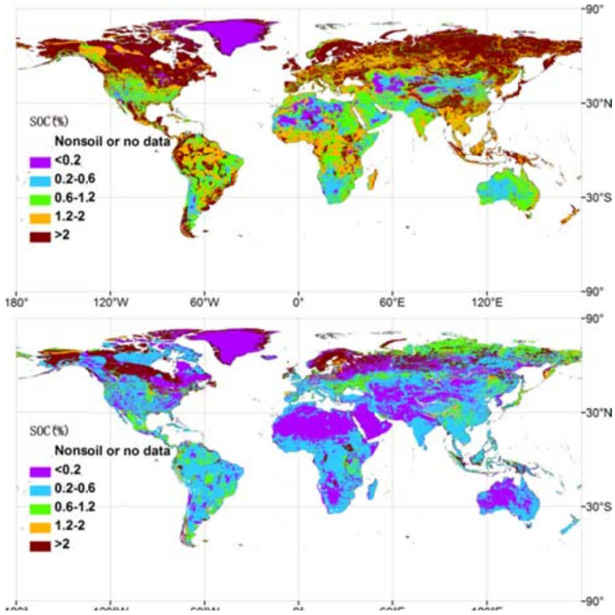

Figure S20. The geographic distribution of organic carbon of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

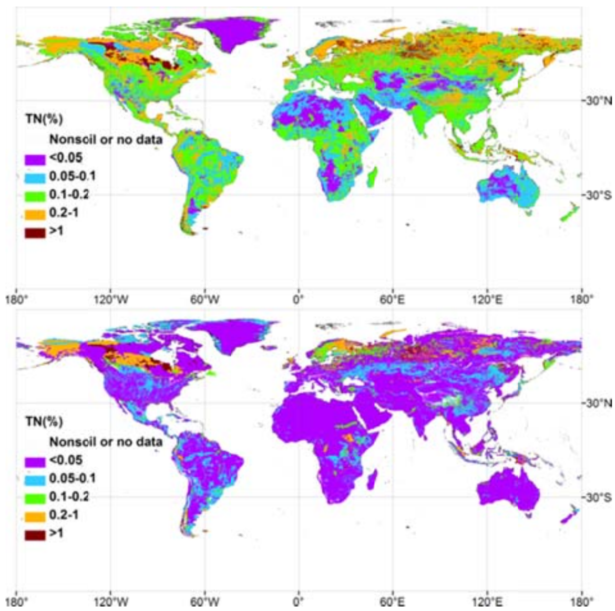

Figure S21. The geographic distribution of total nitrogen of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

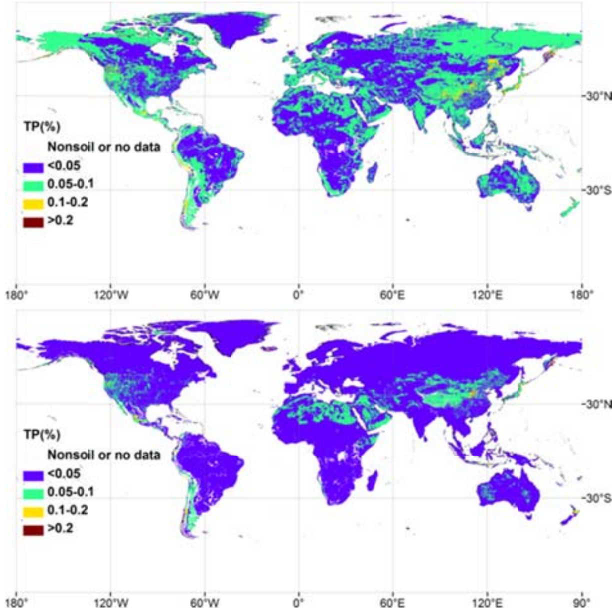

Figure S22. The geographic distribution of total phosphorus of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

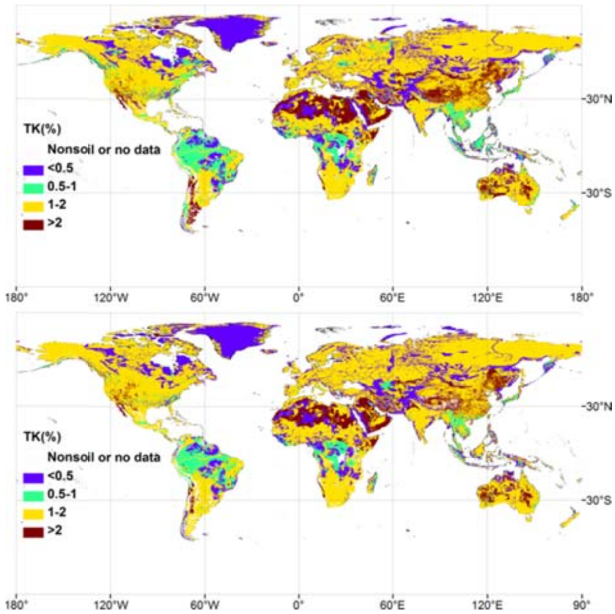

Figure S23. The geographic distribution of total potassium of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

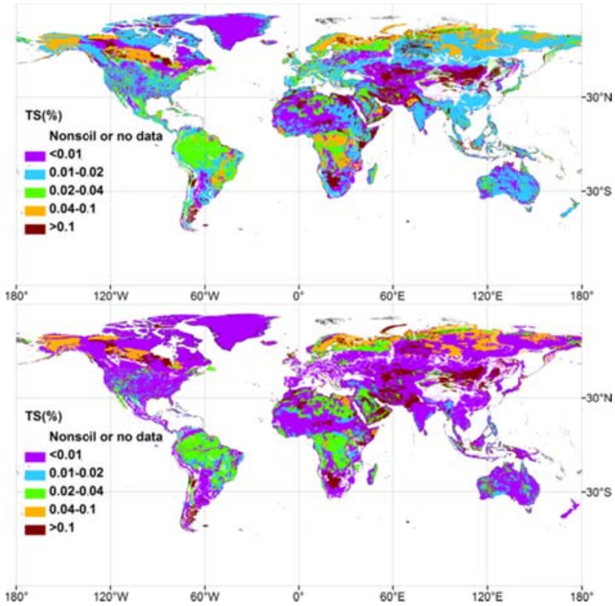

Figure S24. The geographic distribution of total sulfer of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

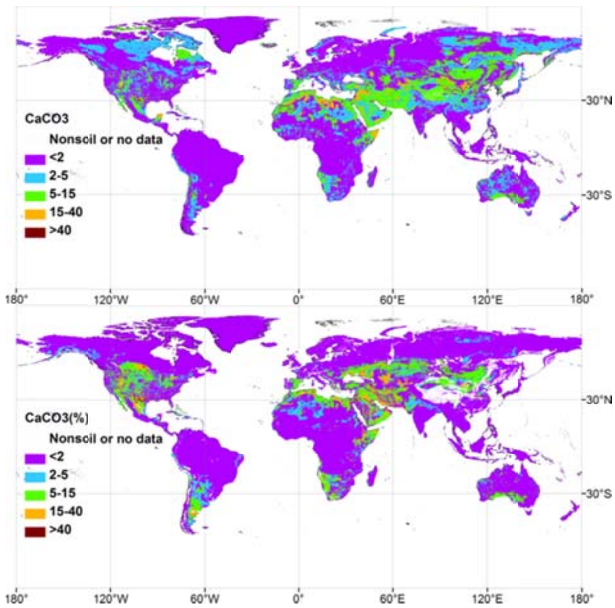

Figure S25. The geographic distribution of calcium carbonate content of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

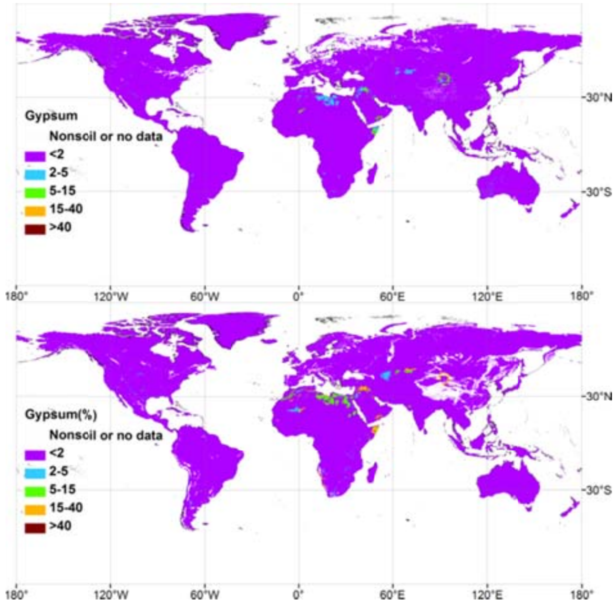

Figure S26. The geographic distribution of gypsum content of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

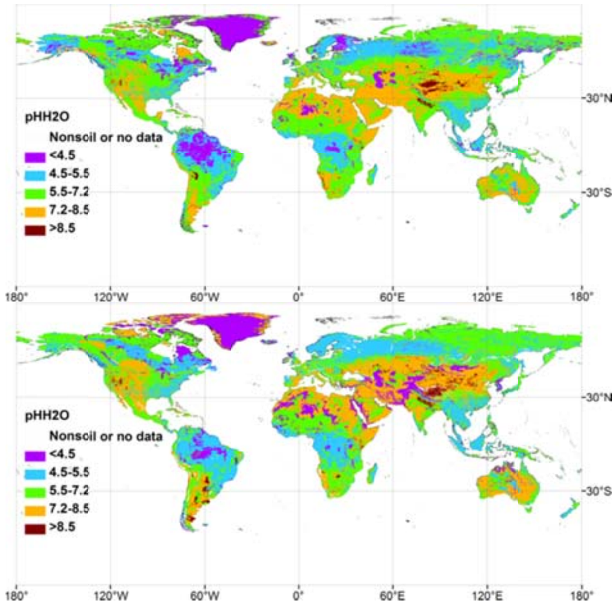

Figure S27. The geographic distribution of pH, measured in water of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

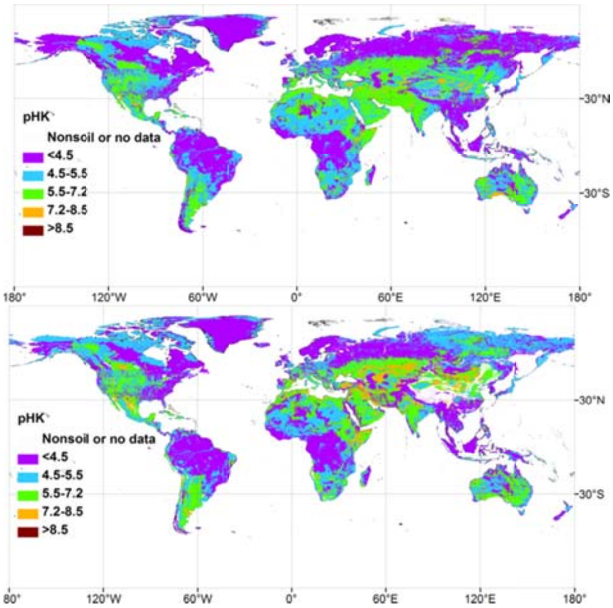

Figure S28. The geographic distribution of pH, measured in KCL solution of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

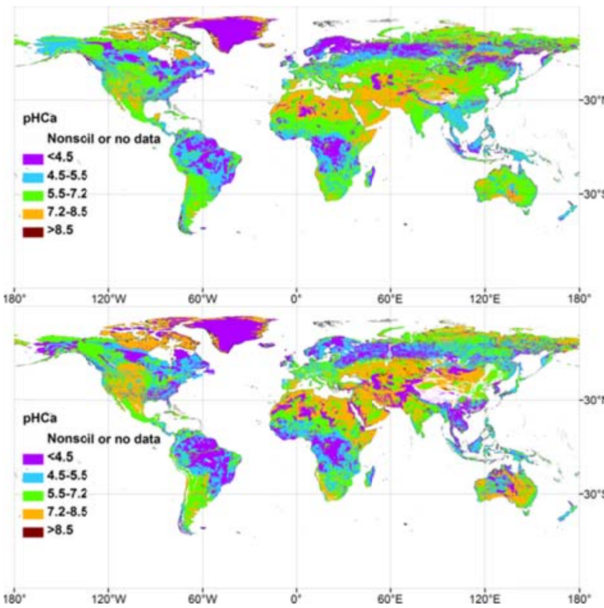

Figure S29. The geographic distribution of pH, measured in CaCl2 solution of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

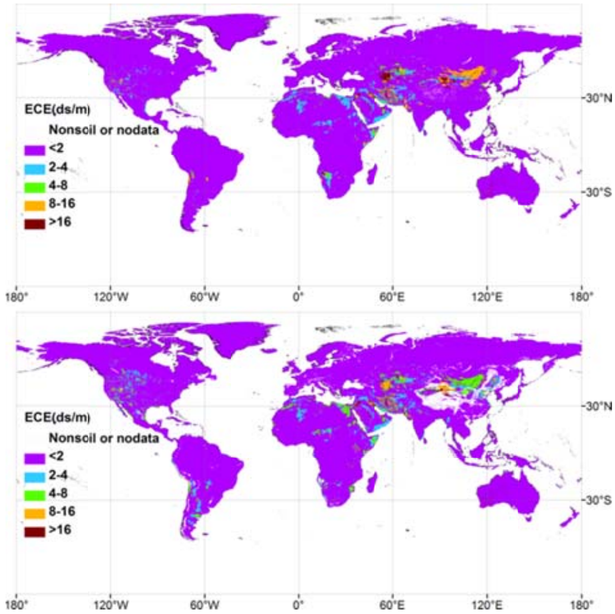

Figure S30. The geographic distribution of electrical conductivity of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

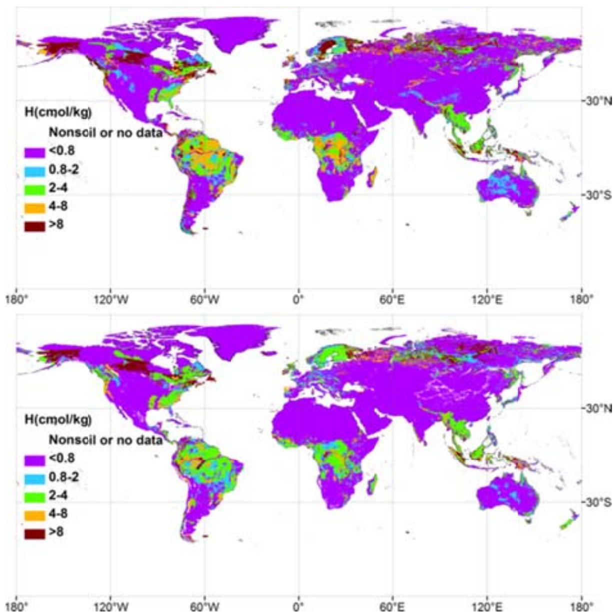

Figure S31. The geographic distribution of exchangeable acidity of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

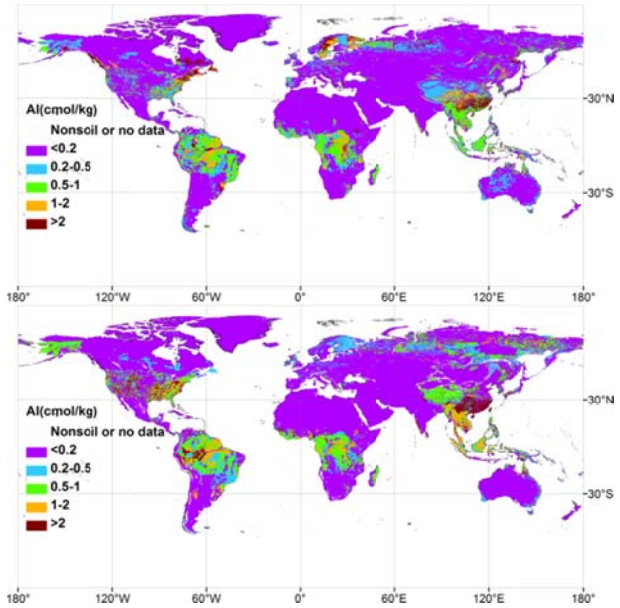

Figure S32. The geographic distribution of exchangeable aluminum of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

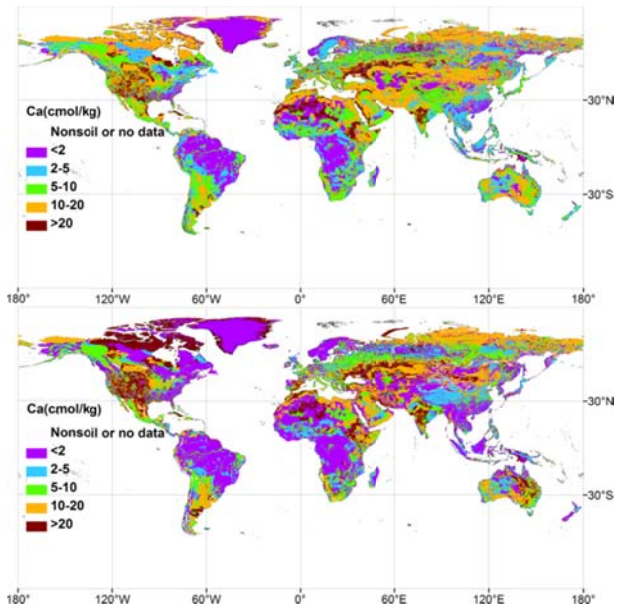

Figure S33. The geographic distribution of exchangeable calcium of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

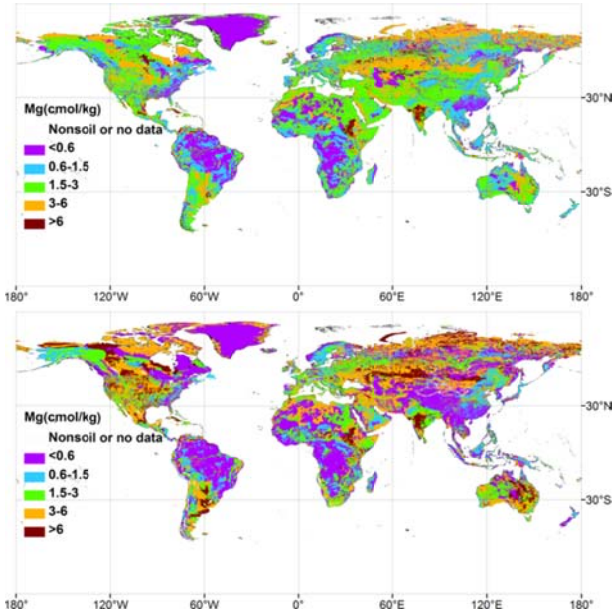

Figure S34. The geographic distribution of exchangeable magnesium of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

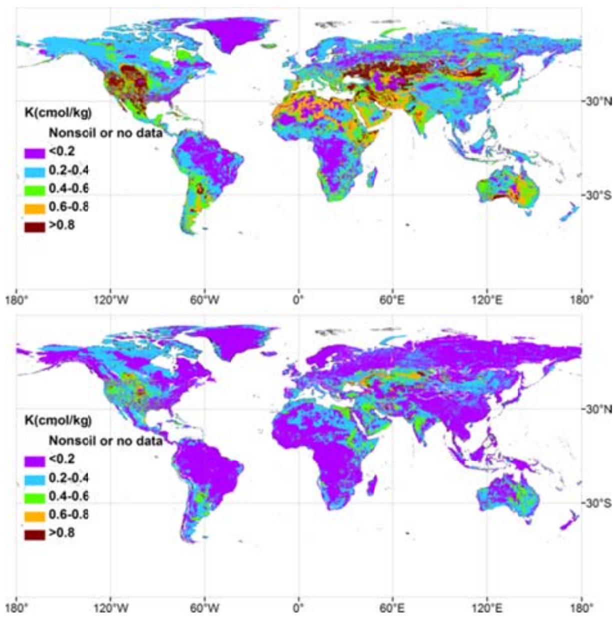

Figure S35. The geographic distribution of exchangeable potassium of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

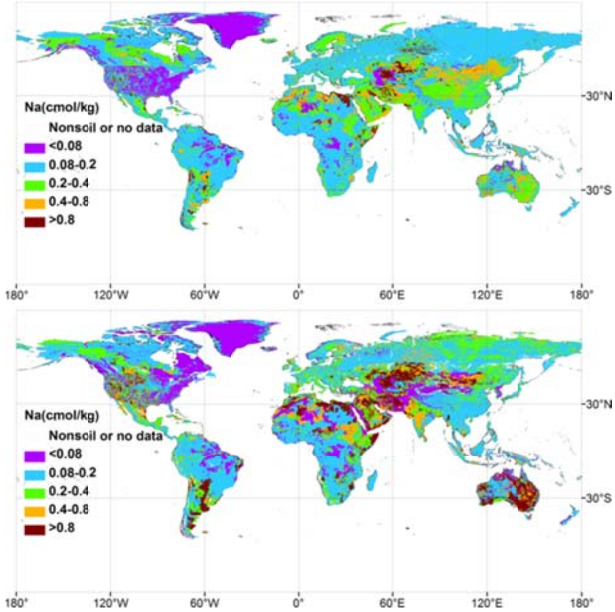

Figure S36. The geographic distribution of exchangeable sodium of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

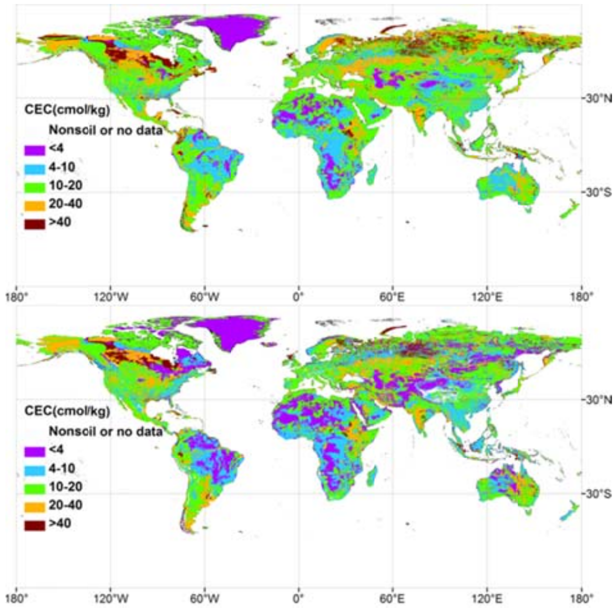

Figure S37. The geographic distribution of cation exchange capacity of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

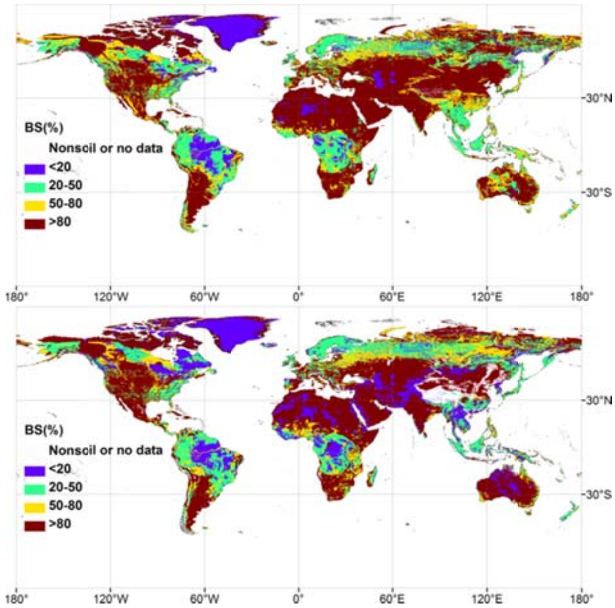

Figure S38. The geographic distribution of base saturation of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

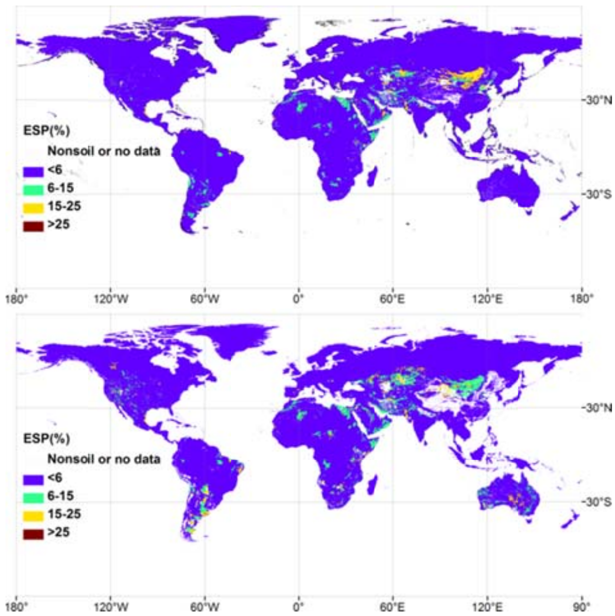

Figure S39. The geographic distribution of exchangeable sodium percentage of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

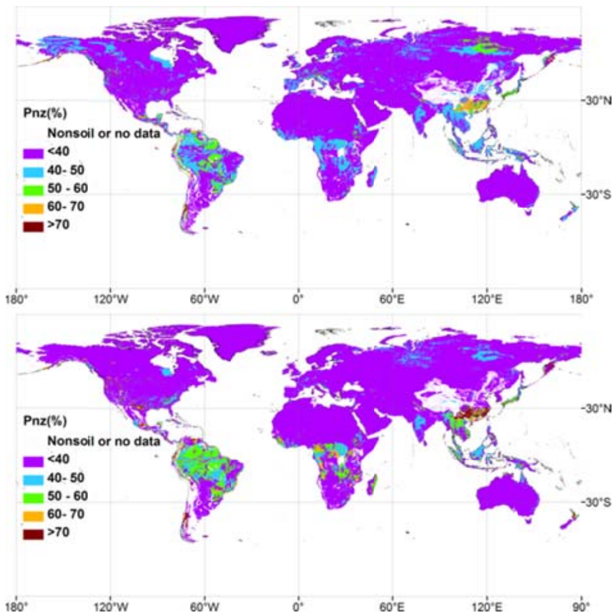

Figure S40. The geographic distribution of phosphorous retention by New Zealand method of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

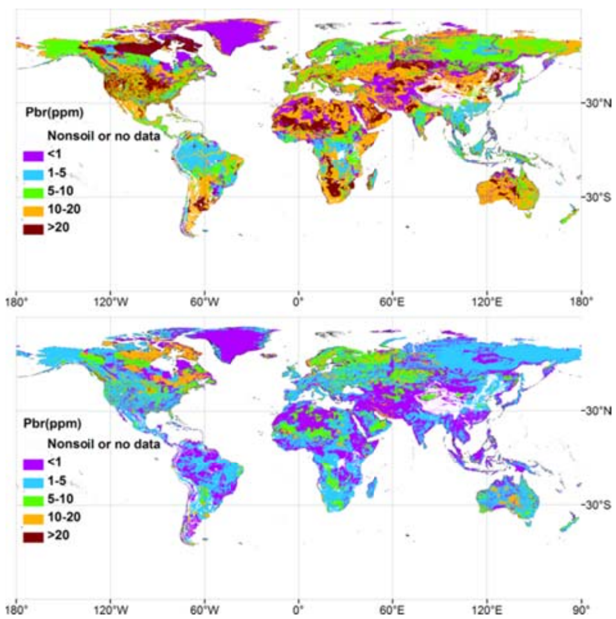

Figure S41. The geographic distribution of phosphorous retention by the Bray1 method of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

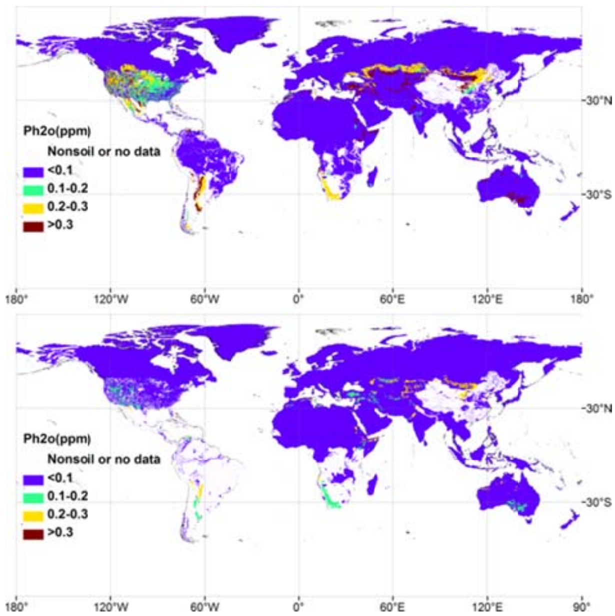

Figure S42. The geographic distribution of the amount of water soluble phosphorous of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

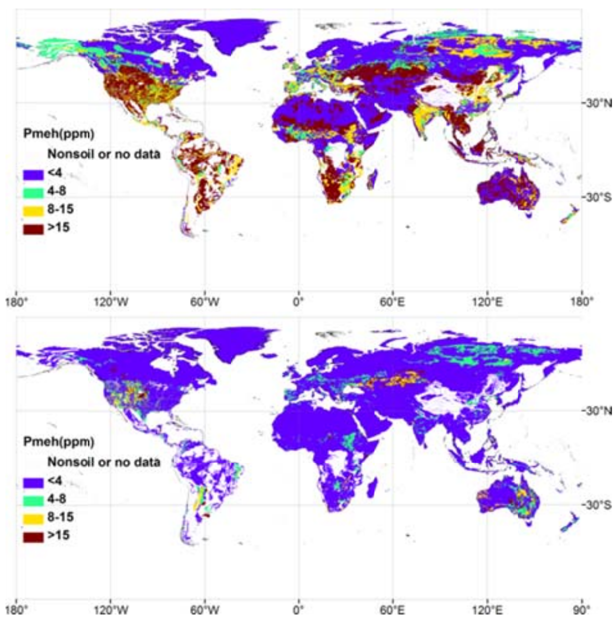

Figure S43. The geographic distribution of phosphorous retention by Mehlich method of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

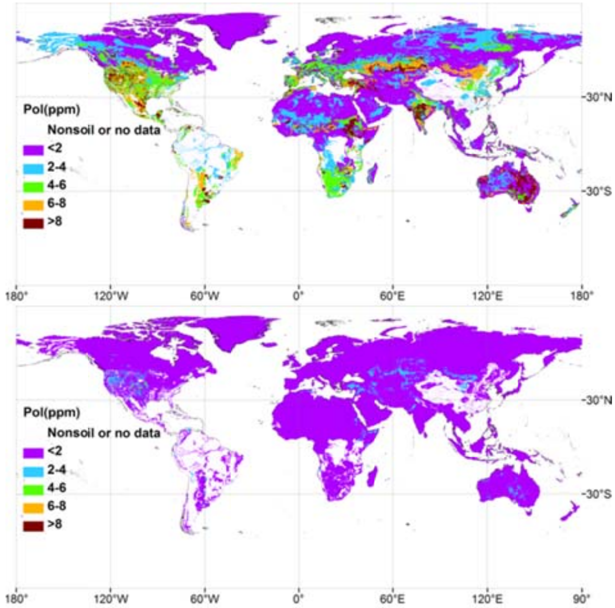

Figure S44. The geographic distribution of phosphorous retention by Olsen method of layer 2 (0.045-0.091 m, upper) and layer 6 (0.493-0.829 m, lower).

Section V. A quick look at data quality

As described in the paper, the soil data quality can be indicated by the quality control information. Here we offer a way for a quick look at data quality. For areas or soil properties where regional soil databases are used, including China, US, part of ESDB, part of SLC and , part of ASRIS, the data quality depends on the regional soil data used. For areas or soil properties where WISE and NCSS soil profiles (Table 1 in the paper), the data quality can be reflected by the abundance of the soil profiles in the following table.

Soil property	Number of profiles	Number of horizons
Total carbon	8158	45151
Organic carbon	31956	179671
Total nitrogen	21089	92998
Total sulfur	5812	33136
Calcium carbonate content	14462	92693
Gypsum content	7831	39797
pH, measured in water	37650	221250
pH, measured in KCL solution	8133	46073
pH, measured in CaCl2 solution	27522	166985
Electrical conductivity	12488	63551
Exchangeable calcium	32170	188502
Exchangeable magnesium	32332	190226

Exchangeable sodium	30841	188005
Exchangeable potassium	32394	191427
Exchangeable aluminum	13386	77835
Exchangeable acidity	5452	26015
Cation exchange capacity	32821	192373
Base saturation, expreesed as % of CEC	27240	166748
Sand content	37025	221351
Silt content	37025	221351
Clay content	37025	221351
Gravel content	21803	139522
Bulk density	11992	81493
Volumetric water content at -10 kPa	13878	11205
Volumetric water content at -33 kPa	9955	72318
Volumetric water content at -1500 kPa	24810	151232
The amount of phosphorous using the Bray1 method	3108	16500
The amount of phosphorous by Olsen method	274	1578
Phosphorous retention by New Zealand method	3294	21548
The amount of water soluble phosphorous	129	569
The amount of phosphorous by Mehlich method	512	3353
Exchangeable sodium percentage	27200	61657
Total phosphorus	974	5291
Total potassium	987	5355